315

[Reprinted from the Journal of the American Chemical Society, 70, 3040 (1948).]

[CONTRIBUTION FROM THE EASTERN REGIONAL RESEARCH LABORATORY[1]]

Water Absorption of Proteins. III. Contribution of the Peptide Group[2]

BY EDWARD F. MELLON, ALFRED H. KORN AND SAM R. HOOVER

Theoretical considerations by several workers[3,4,5] have established a number of correlations between the number of polar groups in certain proteins and the amount of water absorbed by these proteins. Most workers have based these correlations on the possibility that the polar groups form hydrogen bonds with water, and all of them have based the correlation on the total number of the groups present. Recently we[6] have quantitatively determined the contribution of the amino group to the water-absorbing power of casein. This study indicated that the amino group, which constitutes less than 1% of the weight of the casein molecule, is responsible for about one quarter of the total water absorbed by casein over the entire range between 6 and 93% relative humidity. The amount of water held above 75% relative humidity is consistent with the amount required for complete hydrogen bonding of the water to the amino group. It seemed of interest, therefore, to extend the quantitative studies to other polar groups.

The most numerous polar group in proteins is the peptide group, and its contribution to water absorption has been variously reported. Lloyd and Phillips[3] and Pauling[5] have concluded that the hydration of peptide groups would be slight[7]; but Sponsler, Bath and Ellis[4] indicate a maximum hydrogen-bonding capacity of four moles of water per peptide group. They, however, feel that space restrictions and chain interactions would reduce this figure to less than two. Since it is impractical to modify the peptide groups of a protein to reduce their hydrogen-bonding capacity, we have attempted to study a series of peptides with different numbers of peptide groups per molecule. In order to simplify both the synthesis of the peptides and the interpretation of the data, we have limited this study to the glycine peptides. Except for the terminal carboxyl and amino groups, these peptides have no polar groups except the peptide group and there are no side-chain groups to prevent close packing and high chain interaction between the peptide groups.

Experimental

Preparation of Glycine Peptides.—Glycylglycine was prepared from glycine anhydride.[8] Triglycine,[9] tetraglycine[10] and pentaglycine[10] were prepared by adding chloroacetyl chloride to the next lower peptide and then replacing the halogen by ammonolysis. These four peptides were each recrystallized twice from water by the addition of alcohol. Hexaglycine was prepared by alkaline hydrolysis[11] of its methyl ester, which was prepared by

(1) One of the Laboratories of the Bureau of Agricultural and Industrial Chemistry, Agricultural Research Administration, U. S. Department of Agriculture.

(2) This paper is part of a talk presented before the Meeting-in Miniature of the Philadelphia Section of the American Chemical Society, January 22, 1948.

(3) Lloyd and Phillips, *Trans. Faraday Soc.*, **29**, 132 (1933).

(4) Sponsler, Bath and Ellis, *J. Phys. Chem.*, **44**, 996 (1940).

(5) Pauling, THIS JOURNAL, **67**, 555 (1945).

(6) Mellon, Korn and Hoover, *ibid.*, **69**, 827 (1947)

(7) White and Eyring (*Textile Research J.*, **17**, 523 (1947)) have recently analyzed the whole problem of sorption by swelling high polymers. They accepted Pauling's assumptions regarding the contribution of polar side groups and peptide linkages to sorption by proteins

(8) Fischer and Fourneau, *Ber.*, **34**, 2868 (1901)

(9) Fischer, *ibid.*, **36**, 2982 (1903).

(10) Fischer, *ibid.*, **37**, 2486 (1904).

(11) Fischer, *ibid.*, **39**, 453 (1906).

the condensation of triglycine methyl ester.[12] The hexaglycine was recrystallized twice from dilute ammonium hydroxide by evaporation of the ammonia. The precipitate was washed several times with water and once with absolute alcohol. Benzoyltetraglycine was prepared by dissolving 5.5 g. of tetraglycine in 8.5 ml. of normal sodium hydroxide and treating this solution slowly with 3.5 ml. of benzoyl chloride in 50 ml. of dry ether. Additional alkali was added to keep the solution basic. The solution was acidified with hydrochloric acid, and the precipitate was washed with water and 95% alcohol. This crude material was purified by dissolving in alkali, reprecipitating with hydrochloric acid, and washing with water until the supernatant contained no chloride ion. The air-dried material weighed 1.8 g. These peptides were analyzed for nitrogen content by the Kjeldahl method and for amino content by electrometric titration. The analytical values found and calculated are given in Table I.

TABLE I

ANALYSES OF GLYCINE PEPTIDES

Peptide	Nitrogen, %[a] Calcd.	Nitrogen, %[a] Found	NH_2-N, %[a] Calcd.	NH_2-N, %[a] Found
Diglycine	21.2	21.2	10.6	10.6
Triglycine	22.2	21.9	7.3	7.4
Tetraglycine	22.8	22.5	5.6	5.4
Pentaglycine	23.1	23.0	4.6	4.8
Hexaglycine	23.3	23.3	3.8	3.9
Benzoyltetraglycine	16.0	15.9	0.0	0.0

[a] Moisture-free basis.

Polyglycine-I was prepared by the method of Frankel and Katchalski.[13] Free glycine methyl ester (200 g.) was dissolved in 600 ml. of anhydrous ether, and the solution was allowed to stand at room temperature for one-hundred thirty days. The precipitate was filtered off and washed twice with dry ether (yield 120 g.). This dry residue was repeatedly extracted with 500-ml. portions of boiling water until biuret and picric acid tests on the filtrate indicated that all soluble peptides and glycine anhydride had been removed (about 8 times). The material was then filtered and dried over calcium chloride (yield 17 g.). It contained 23.10% nitrogen and 1.23% methoxyl, corresponding to 41 moles of nitrogen to one mole of methoxyl and indicating an average chain length of 41 glycine residues. The observed value for nitrogen however, does not satisfy the requirements for a polyglycine methyl ester of this chain length, and it is possible that the Kjeldahl nitrogen analysis did not determine all the nitrogen present. On the basis of the methoxyl content alone, the average peptide length should be 44 units long, and the nitrogen content should be 24.3%. The insolubility of the materials and the spurious results given by glycine peptides in the Van Slyke amino nitrogen determination make it impossible to define the chain length of these materials more accurately. We do know that there were no peptides shorter than hexaglycine and that no glycine anhydride was present.

Polyglycine-II was prepared by the method of Pacsu and Wilson.[12] Triglycine methyl ester hydrochloride (125 g.) was converted to the free ester with sodium methylate. The free ester was isolated, and the crystalline material was heated at 101 ± 1° for sixty-two hours. The dry esters (57 g.) were extracted with six 100-ml. portions of hot methyl alcohol and ten 100-ml. portions of hot water. The product was dried by sublimation of the water, giving 35 g. of the polyglycine methyl ester. This material contained 23.33% nitrogen and 0.91% methoxyl, corresponding to 54 moles of nitrogen and one mole of methoxyl and indicating an average chain length of 54 glycine residues. The observed value for nitrogen does not satisfy the requirement for a polypeptide ester of this chain length, probably for the same reason given for the polyglycine-I. On the basis of the methoxyl content alone, the average peptide length will be 59 units long and the nitrogen content should be 24.4%.

The Absorption Determination.—About 2-g. quantities of the compounds to be studied were dried under a 29-in. vacuum at 70° with a stream of dry air flowing through the oven to remove the liberated moisture.[6] The samples were dry within fourteen hours. A subsequent drying period of six hours demonstrated that an equilibrium dry weight had been attained. The dried samples were put into desiccators over saturated salt solutions, and after evacuation to about 40 mm. pressure the desiccators were submerged in a water-bath the temperature of which was regulated to 30.0 ± 0.1°. The samples were removed from the desiccators and weighed every three days until equilibrium was demonstrated, after which they were removed to a higher humidity and the equilibration procedure was repeated. After the humidity studies were completed on a sample, it was redried to demonstrate the constancy of the dry weight. The salt solutions used and other details of the procedure are described fully in the first paper of the series.[6]

Results and Discussion

Glycine, glycylglycine and triglycine have shown a complete lack of absorption of water vapor at all humidities up to 93% relative humidity. This may be due to the complete utilization of the hydrogen-bonding capacity of these molecules in producing their crystal structure. Frey and Moore[14] have recently reported adsorption studies on glycine and leucine crystals. The magnitude of their adsorption, which appears to be a surface phenomenon, is too small to be measured by our weighing technique. Tetraglycine, however, which gives a crystalline X-ray diagram[15] similar to that of the di- and tripeptides, shows (curve 2, Fig. 1) a considerable absorption of water from the vapor phase. In order to determine whether this absorption was due to the terminal amino groups in tetraglycine, we have made similar absorption studies on benzoyltetraglycine, which contains no free amino groups. This material has a definite absorption (curve 1, Fig. 1), but it is considerably less than the absorption of tetraglycine, indicating that the terminal amino group is responsible for some of the water absorption but that the carboxyl or peptide groups also show some affinity for water.

Penta- and hexa-glycine, which are not so definitely crystalline as the shorter peptides but give X-ray diagrams between those of the highly crystalline glycine and the amorphous polyglycines,[15] absorb even more water vapor than the tetraglycine. The curve for hexaglycine is given in Fig. 1, and the curve for pentaglycine would be between those for the tetra and hexa polymers. The fact that these glycine polymers absorb more water vapor as the number of glycine residues increases seems to be analogous to the increase of dielectric increment of these compounds with increase in the number of glycine residues demonstrated by Wyman and McMeekin.[16] The ionization constants

(12) Pacsu and Wilson, *J. Org. Chem.*, **7**, 117 (1942).

(13) Frankel and Katchalski, THIS JOURNAL, **64**, 2264 (1942).

(14) Frey and Moore, Abstracts, 112th Meeting, American Chemical Society, September, 1947.

(15) Meyer and Go, *Helv. Chim. Acta*, **17**, 1488-1492 (1934).

(16) Wyman and McMeekin, THIS JOURNAL, **55**, 908 (1933).

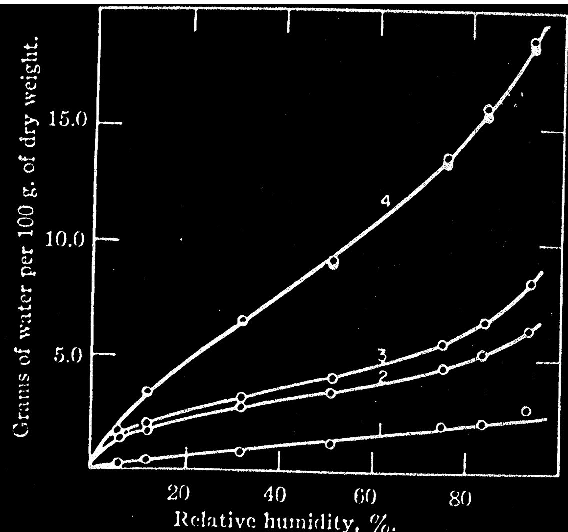

Fig. 1.—Sorption of glycine polymers: 1, benzoyltetraglycine; 2, tetraglycine; 3, hexaglycine; 4, polyglycine esters.

for both the carboxyl and amino groups are practically constant for the four-, five- and six-unit peptides,[17] and consequently we feel that the hydrogen-bonding capacity of these end-groups probably would also remain fairly constant. The amount of water absorbed by these groups would then be a function of the fraction of the total hydrogen-bonding capacity of these groups not required to maintain the crystal structure. This fraction has not been determined, and we are unable to show definitely whether the increased absorption with increase in peptide length is due entirely to the additional peptide groups or whether it is due in part to the end-groups of the molecule.

The two non-crystalline[16] polyglycine esters show almost identical absorption isotherms (Curve 4, Fig. 1), although they have considerably different chain lengths. The amount of water absorbed is much greater than that absorbed by hexaglycine, and on a molar basis (Table II) is much more than can be accounted for by the amino and ester chain ends of the molecule. The peptide groups must be responsible, therefore, for most of the water absorbed by these polyglycine esters. These materials are polypeptide chains similar to the backbone chains of the proteins, and they differ from the natural peptide chains only in their shortness and lack of side groups from the alpha carbons. We conclude, therefore, that the peptide chain of proteins is responsible for a large portion of the vapor-phase water absorption of proteins. In an attempt to determine the magnitude of this effect, we have compared (Table III) the water absorption of several proteins at 60% relative humidity. The nitrogen content of these materials has been divided according to Chibnall[18] into two parts—the nitrogen in the backbone chain and the nitrogen in the side groups. The number of peptide groups in 100 g. of the protein is proportional to the percentage of nitrogen in the backbone chain. The water absorbed per gram of nitrogen in the chain, therefore, expresses the relation of the total absorption (regardless of site) to the number of peptide groups present in the various proteins. Ellis and Bath[19] have shown by infrared absorption spectra that the imino groups of the peptide bond are involved in the absorption of water by gelatin, and Dole[20] has recently shown that in certain of the elastic nylons, where all the hydrogens of the imino nitrogens have been replaced by alkyl radicals, there is still an appreciable water absorption which presumably is due to the carbonyl groups. Our data do not enable us to distinguish between the absorptions of the imino and carbonyl groups, and we do not wish to imply by the basis of our comparison that only the imino nitrogen is involved but rather that the nitrogen content of the chain is a measure of the number of peptide groups.

Table II

Water Sorption per Mole by Glycine Peptides

Relative humidity, %	Water absorbed, moles/mole of peptide: 4^a	$4B^b$	5^a	6^a	44^c	59^d
5.9	0.19	0.06	0.23	0.34	..	..
11.8	.24	.08	.28	.41	4.9	6.4
31.4	.38	.16	.45	.63	9.2	12.4
50.9	.48	.25	.58	.82	12.8	17.0
75.1	.63	.40	.78	1.14	19.2	25.4
83.6	.72	.44	.87	1.32	22.1	29.2
93.3	.88	.57	1.04	1.67	26.2	35.0

[a] Number of glycine residues per molecule. [b] Benzoyl tetraglycine. [c] Polyglycine I. Molecular weight is 2500. [d] Polyglycine II. Molecular weight is 3400.

Table III

Sorption and Peptide Nitrogen

Protein	N in chain, %	N in side groups, %	Sorption at 60% R. H.: %[a]	g./g. N (chain)
Nylon	12.3	0.00	4.1	0.33
Silk	18.5	0.22	8.9	.48
Polyglycine	24.4	0.00	10.9	.45
Zein	12.6	3.53	8.0	.64
Casein	12.2	3.56	11.8	.97
Benzoyl casein	12.2[b]	3.56[b]	9.5[b]	.78

[a] g./100 g. dry weight. [b] These values are calculated on the basis of 100 g. of dry casein which has its lysine epsilon amino groups benzoylated and thus not available for absorbing water.

Unoriented nylon, which resembles the proteins and polyglycine peptides in that it contains peptide-like linkages, shows an absorption per peptide group which is only about 70% of the value shown by silk and the polyglycines. This difference, however, could easily be explained by the presence of 30% of the material in a crystalline state. Crys-

(17) Cohn and Edsall, "Proteins, Amino Acids and Peptides," Reinhold Publishing Corp., New York, N. Y., 1943, p. 84.

(18) Chibnall, *Proc. Roy. Soc.* (*London*), **B131**, 136 (1942).

(19) Ellis and Bath, *J. Chem. Phys.*, **6**, 723 (1938).

(20) Dole, Abstracts, 112th Meeting, American Chemical Society, September, 1947.

talline nylon may differ greatly from crystalline proteins because of the length of the hydrocarbon chain between the peptide groups. Bunn and Garner[21] have recently described the crystal structure of nylon as a series of sheets of molecules. The individual chains are held together by hydrogen bonding in the sheets, and only van der Waals forces appear to be active between the sheets. Such a closely packed structure should make the peptide bonds quite unavailable to water vapor, and it would appear that even the most highly crystallized nylons must have considerable amorphous regions to explain their water absorption.

Silk, which is a protein containing a predominance (75%) of the simple amino acids, glycine and alanine, and only a small number of polar side chains, has an absorption per peptide link almost identical with that shown by our polyglycine molecules. Most other proteins show a somewhat higher value, owing to their polar groups and possibly to the more open structure produced by their side chains.[22] Senti[23] has given a good representation of the manner in which side chains modify the structure and packing of polypeptide chains. Zein[24] and casein, both of which contain a relatively large number of polar side chains, show correspondingly higher water absorptions per peptide group. If the absorption due to the epsilon amino groups of lysine is subtracted from the absorption by casein, the value obtained is much nearer that for zein, which contains only a trace of lysine.

If we consider the 0.45 g. of water per gram of chain nitrogen absorbed by the polyglycine esters (Table III) as the absorption of a bare peptide chain (including the chain ends) at 60% relative humidity, then the peptide groups in casein would be responsible for about 46% of the total water absorbed by casein. This contribution of the peptide linkage is a relatively constant proportion of the total sorption throughout the whole range of relative humidity, 45 ± 3%. The comparison would therefore be the same at other humidities; 60% R.H. was chosen for convenience. Since we have previously attributed[6] about 25% of the absorption by casein to the amino group, only about 30% of the vapor-phase water absorption by casein remains unaccounted for. By a similar calculation, about 70% of the vapor-phase water absorption of zein can be attributed to the peptide groups present.

Since the vapor-phase water absorption of polyglycine polymers can be used as above to approximate the water held by the peptide chain in the proteins and other polypeptides, it is desirable to be able to interpolate and extrapolate our experi-

(21) Bunn and Garner, *Proc. Roy. Soc.* (*London*), A189, 39-68 (1947).

(22) We have shown experimentally that the degree of crystallinity of silk and ovalbumin does not affect the water absorption markedly. These results are being prepared for publication.

(23) Senti, *Am. Dyestuff Reporter*, 36, 230 (1947).

(24) 25° data of Bull, This Journal, 66, 1499 (1944).

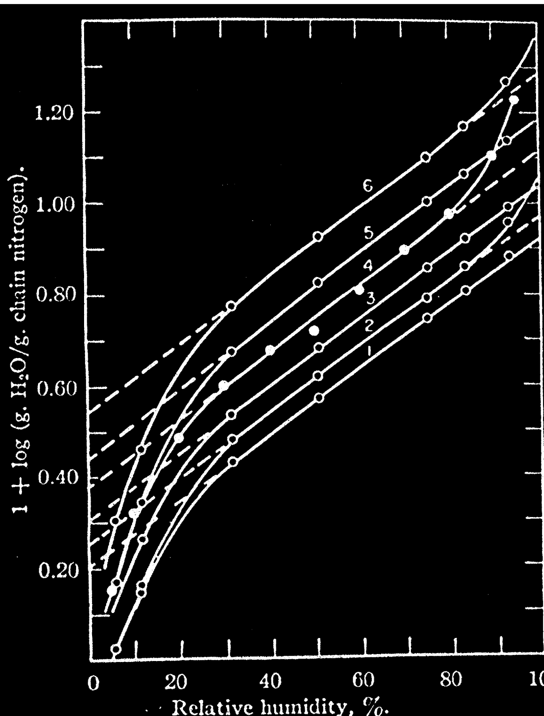

Fig. 2.—Water absorption of peptides and proteins at 30°: 1, polyglycine ester II; 2, silk fibroin; 3, polyglycine ester I (ordinate increased by 0.1 units); 4, zein (Bull's 25° data); 5, ε-amino-benzoylated casein; 6, casein.

mental values to give accurate values at other relative humidities. The empirical plot shown in Fig. 2 was developed by graphical analysis of our data. It is extremely useful in interpolating absorption data at humidities from 30 to 80%. When the logarithm of the absorption value is plotted against the relative humidity, a straight line is obtained between 30 and 80% relative humidity for most proteins. With some proteins and peptides, the linear region appears to extend beyond 93% relative humidity, and, therefore, may give a reasonable extrapolation to 100% relative humidity. Our two polyglycine molecules and our ε-amino benzoylated casein give curves capable of this extrapolation to 100% relative humidity. Silk fibroin and casein show some deviation from the straight line above 90% relative humidity, but a curvilinear extrapolation to 100% relative humidity can be made with only a moderate increase in the error of the result. This system of extrapolation appears to be superior to that used by Bull.[24]

The curves of Fig. 2 have been plotted with the absorption shown in grams of water per gram of nitrogen in the backbone chain as a convenience for further calculation. Any unit for measuring the absorption and any measure of the vapor pressure of water present can be used to give this type of curve. The equation of the straight line portion of the curves can readily be calculated. The slopes and intercepts for the various proteins are given in Table IV.

Table IV

Constants[a] for Equation Log $R = AH + B$ and the Absorption Value (R) in Saturated Water Vapor

Protein	A	$B + 1$	R^a at saturation
Polyglycine I	0.00736	0.202	0.87
Polyglycine II	.00720	.201	0.83
Silk	.00723	.251	1.08[b]
Zein	.00736	.376	
Casein	.00743	.540	2.46[b]
ϵ-Amino free casein	.00748	.438	1.54

[a] R = g.-H_2O/g.-N in chain; H = % relative humidity. [b] By curvilinear extrapolation.

The extrapolated intercepts at zero relative humidity are probably related to absorption of water at low humidities, which is more firmly held than that absorbed at higher humidities. These intercepts are strictly empirical, for the logarithmic function approaches minus infinity as the sorption approaches zero. The intercept does increase with increase in the number of polar groups in the side chains of the proteins, but the correlation at present does not seem to be quantitative. This is to be expected, since all polar groups do not have the same type of absorption, and therefore the sum of their absorption values will depend more on the relative proportion of each polar group than on the total number of polar groups. This indicates the need for detailed knowledge of the water-absorbing properties of each polar group in the proteins.

The slopes of these lines are remarkably constant when we consider that a number of different groups must contribute to the total water absorption value. This fact indicates that all the water absorbed in this humidity region is held by hydrogen bonds of about the same strength. The line for nylon (not shown) has a much greater slope than that for the proteins, thus emphasizing the difference between nylon and the polypeptides already shown in Table III and the previous discussion.

Table V shows the B.E.T. constants[25] calculated from the data for the polyglycines and proteins discussed above. The value for V_m, which corresponds to the amount absorbed in a monolayer, is considerably less than one mole of water per mole of peptide nitrogen even though we know this value includes the water which is held on other polar groups of the molecule.

The correlation between the amount absorbed in a monolayer as calculated by the B.E.T. treatment and the extrapolated intercept to zero relative humidity given by our isotherm is remarkable. The intercept value is about three quarters of the B.E.T. monolayer value in each case, and our extrapolated isotherms give values equivalent to the B.E.T. monolayer at about 15-19% R.H. This humidity is just slightly lower than the humidity at which the materials have absorbed water equivalent to the B.E.T. monolayer.

(25) Brunauer, Emmett and Teller, This Journal, 60, 309 (1938).

Table V

B. E. T. Constants Compared with the Intercepts of the Semilog Plots

Protein	B. E. T. constants V_m[a]	B. E. T. constants $E_1 - E_L$[b]	Intercepts of semilog straight lines[a] 0% R. H.	Intercepts of semilog straight lines[a] 100% R. H.	0% R. H. intercept/V_m
Polyglycine I	0.160	1.55	0.124	0.678	0.78
Polyglycine II	.159	1.51	.124	.656	.78
Silk fibroin	.184	1.38	.139	.739	.76
Zein	.247	1.53	.185	1.01	.75
ϵ-Aminobenzoylated casein	.297	1.27	.214	1.20	.72
Casein	.363	1.37	.270	1.49	.74

[a] Moles of water per mole of peptide nitrogen. [b] Kcal. per mole.

None of the previously developed theories applies to our empirical isotherm, but the applicability of this isotherm and its physical significance are being investigated.

Summary

The vapor-phase water absorption of glycine peptides from two to six units in length has been determined, and the values show clearly that non-hygroscopic amino acids may give hygroscopic peptides.

Polyglycine peptides of longer chain length show that the peptide linkage must be responsible for most of the water absorption by these materials.

Comparison of the absorption of the polyglycine peptides with the absorption of proteins indicates that the absorption of the peptide chain backbone is probably of the same magnitude, if not identical, for all long-chain polypeptides and proteins.

Peptide groups appear to be responsible for about 45% of the vapor-phase water absorption by casein and 70% of the absorption by zein at 60% relative humidity.

A new absorption isotherm for the high humidity range has been presented which permits a linear interpolation between 30 and 80% relative humidity for all proteins and a further linear extrapolation to 100% relative humidity in a few special cases.

Philadelphia 18, Pa. Received February 19, 1948